

PAT-NO: JP411088847A
DOCUMENT-IDENTIFIER: JP 11088847 A
TITLE: VIDEO/AUDIO SYNCHRONIZING SYSTEM
PUBN-DATE: March 30, 1999

INVENTOR-INFORMATION:
NAME
USUI, SHUJI

INT-CL (IPC): H04N007/06, H04N005/00 , H04N007/24

ABSTRACT:

PROBLEM TO BE SOLVED: To automatically correct the sense of incompatibility caused by a timing difference in the transmission delay of image and sound by detecting the timing error of video and audio signals and synchronizing both the video and audio signals by delaying one of these video and audio signals at the timing when the other signal is delayed.

SOLUTION: A video signal feature extractor 18 extracts a signal component corresponding to predetermined operation from a video signal (a) and outputs video timing (g). An output signal (c) of a decoder 12 for audio signal is applied to an audio signal feature extractor 22 and audio timing (h) is outputted. The video timing (g) and audio timing (h) are applied to a time error detector 20 and an image display 23. The time error detector 20 detects the difference from the respective video and audio timing (g) and (h) and applies error time (i) of both the timing to a delay circuit 21. As a result, feedback control is performed so that the timing of decoder 10 for video signal and decoder 12 for audio signal can be zero, and the feature timing of video and audio signals is matched.

COPYRIGHT: (C) 1999, JPO

----- KWIC -----

【特許請求の範囲】

【請求項1】 画像処理によって映像信号から拍手や口の動き等に相当する所定条件のゼスチャ部分の発生タイミングを検出する手段と、音声信号から上記映像信号のゼスチャ部分に対応する所定条件部分の発生タイミングを検出する手段を有し、両者のタイミング誤差を検出するとともに、映像と音声のどちらか遅れたタイミングに他方を遅延させることにより、映像信号と音声信号の同期をとることを特徴とする映像・音声同期方式。

【請求項2】 請求項1の映像・音声同期方式において、音声信号の所定条件部分の発生タイミングと映像信号の所定条件部分の発生タイミングを同一時間軸上に表示し、映像信号と音声信号の同期確認を容易にすることを特徴とする映像・音声同期方式。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、映像信号と音声信号を異なる伝送媒体を経由して伝送した時に、受信側で発生する映像信号と音声信号の時間差を合致させるシステムに関するものである。

【0002】

【従来の技術】近年、動画像圧縮の国際標準(MPEG2方式)に準拠した映像信号や音声信号の高効率符号化が実使用に供されている。また、圧縮されたデジタルデータは、有線のみならず、光ケーブルや衛星等様々な伝送媒体を経由して伝達されている。一般に、映像信号と音声信号は同一伝送媒体で伝送するため、受信側における映像信号と音声信号の時間的タイミング誤差は、エンコード/デコードに伴う処理遅延とデータ多重部や多重分離部で原理的に発生する遅延量であるため、予測可能な遅延量である。このタイミング誤差に対し、MPEG2のシステムでは、タイムスタンプと呼ばれる時間情報を映像と音声の圧縮データにそれぞれ多重し、受信側でのタイミング合わせとして用いる。しかし、映像信号と音声信号を別々の回線を経由して伝送する場合には、受信側における遅延時間の補償量の設定が困難な状況となる。

【0003】図2は、異なる伝送媒体を経由して映像信号と音声信号を伝送するシステムの一例である。カメラ1等から得られた映像信号aは、映像信号用高能率符号器2で圧縮され、圧縮映像データbとなる。一方、マイク3等から得られる音声信号cも音声信号用高能率符号器4で圧縮され、圧縮音声データdとなる。圧縮映像データbと圧縮音声データdは、データ多重器5でそれぞれタイムスタンプを付され、一束のシリアルデータeとなり、衛星送信機6、静止衛星7を経由して衛星受信機8に到達する。衛星受信機8の出力シリアルデータfは、誤りが無ければ、シリアルデータeと同一である。この後、送信側と逆の処理が施され、多重分離器9で圧縮映像データbと圧縮音声データdに分離され、映像信

号用デコーダ10で映像信号aに復元され映像信号モニタ11で表示される。一方、圧縮音声データdも音声信号用デコーダ12で元の音声信号cに復元され、スピーカ13を駆動する。

【0004】以上説明したごとく、一般的な伝送形態では映像信号と音声信号とは、同一の衛星回線で送られる。しかし、衛星回線は光ケーブル伝送に比べ、大気や気象条件の影響を受けやすく、光ケーブル伝送等に比べ本質的に信頼度が低い。このため、番組編成上、映像信号は最悪途切れても、音声信号だけは確保したいとの考えから、映像信号は衛星回線で、音声信号は光ケーブル回線で運用することが多い。このような音声信号のみを光ケーブル14で伝送する例を図2に示す。圧縮音声データdは、光ケーブル用送信インターフェース15で光ケーブル14に最適変調を施され、光ケーブル用受信インターフェース16で再び圧縮音声データdに変換され、音声信号用デコーダ12に到達する。この時間問題になるのが、映像信号と音声信号のタイミングである。一般に、違和感のない映像信号と音声信号のタイミング量は、映像信号に比べ音声信号が、進みで約30ミリ秒、遅れで約60ミリ秒とされている。ここで、具体的な遅延量を要素別に考える。初めに、原理的に発生する遅延として、映像信号は高能率符号化/復号化処理により、約500ミリ秒の遅れが発生する。一方、音声信号処理は映像信号処理に比べ短時間に処理でき、その時間は数十ミリ秒である。

【0005】この結果、通常の処理では400ミリ秒以上のタイミング誤差が発生するため映像信号を遅らせなければならないことは明白である。さらに、実用に供されるシステムで、静止衛星を経由すると36000kmの往復電波伝搬で、約250ミリ秒が必要となる。実例として、地球の裏側から日本への伝送では、最悪4回の静止衛星中継が行われ、約1秒の遅延が固定的に加わる場合も珍しくない。このため、伝送媒体に起因するタイミング誤差は、通常、遅れの少ない音声信号を遅延の大きな映像信号に合うように遅延させることで運用しているが、大まかな設定をした後は、人間の感覚で微調整していた。

【0006】

【発明が解決しようとする課題】前述のごとく、従来技術では、最終遅延調整は人間に頼っていたため、多地点からの生中継で番組を構成する場合、遅延時間の調整が困難であった。本発明はこれらの欠点を除去し、リップシンクと呼ばれる映像と音声の伝送遅延のタイミング差に起因する違和感を自動的に補正することを目的としている。

【0007】

【課題を解決するための手段】本発明は、上記の目的を達成するため、画像処理によって映像信号から拍手や口の動き等に相当する所定条件のゼスチャ部分の発生タイ

3

ミングを検出する手段と、音声信号から上記映像信号のゼスチャ部分に対応する所定条件部分の発生タイミングを検出する手段を有し、両者のタイミング誤差を検出するとともに、映像と音声のどちらか遅れたタイミングに他方を遅延させることにより、映像信号と音声信号の同期をとるものである。さらに、音声信号の所定条件部分の発生タイミングと映像信号の所定条件部分の発生タイミングを同一時間軸上に表示し、映像信号と音声信号の同期確認を容易にすることを特徴とするものである。

【0008】

【発明の実施の形態】以下、本発明の一実施例を図1を用いて詳細に説明する。なお、前述と同一ブロックには同一番号を付し説明を省略する。本発明は、送信側には一切手を加えず、受信側のみの処理で対応可能である。図1において、映像信号用デコーダ10で復元された映像信号aは、映像信号特徴抽出器18と信号加算器19に加えられる。映像信号特徴抽出器18は、映像信号の中から拍手やアナウンサーの口の動きなど予め決められた動作に相当する信号部分を検出し、映像タイミングgを出力する。この映像タイミングgは、時間誤差検出器20と画像表示器23に加えられる。

【0009】一方、音声信号用デコーダ12の入力側には遅延回路21が追加され、圧縮音声データdが遅延回路21を経由して音声信号用デコーダ12に加えられる。音声信号用デコーダ12の出力音声信号cは、音声信号特徴抽出器22に加えられる。音声信号の中から拍手や人間の音声における破裂音等の特徴に相当する信号部分を検出し、音声タイミングhを出力する。この音声タイミングhは、時間誤差検出器20と画像表示器23に加えられる。時間誤差検出器20では、映像タイミン

4

グgと音声タイミングhから、差分を検出し、両者の誤差時間iを遅延回路21に加える。この結果、映像信号用デコーダ10と音声信号用デコーダ12のタイミングがゼロになるようフィードバック制御され、映像信号と音声信号の特徴タイミングが合致する。本発明ではさらに、映像タイミングgと音声タイミングhを画像表示器23で同一時間軸上に表示する信号を出力し、信号加算器19で映像信号に多重して表示する。

【0010】

10 【発明の効果】以上説明した如く、本発明によれば、人手に頼っていた映像信号と音声信号のタイミング合わせが自動化されると共に、タイミング誤差を画像モニタ上に表示することで、信頼度の高い映像・音声同期システムが構築出来る。

【図面の簡単な説明】

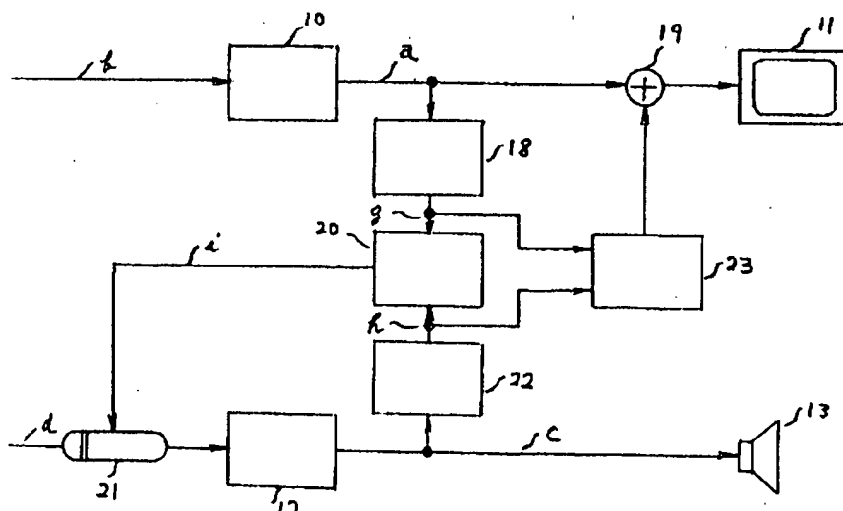
【図1】本発明の全体構成を示すブロック図

【図2】従来のシステムの一例を示すブロック図

【符号の説明】

1：カメラ、2：映像信号用高能率符号器、3：マイク、4：音声信号用高能率符号器、5：データ多重器、6：衛星送信機、7：静止衛星、8：衛星受信機、9：多重分離器、10：映像信号用デコーダ、11：映像信号モニタ、12：音声信号用デコーダ、13：スピーカ、14：光ケーブル、18：映像信号特徴抽出器、19：信号加算器、20：時間誤差検出器、21：遅延回路、22：音声信号特徴抽出器、23：画像表示器、a：映像信号、b：圧縮映像データ、c：音声信号、d：圧縮音声データ、g：映像タイミング、h：音声タイミング、i：誤差時間

【図1】



【図2】

